Chapter 19

A Metaanalysis of Research on Iodine and Its Relationship to Cognitive Development

Nico Bleichrodt and Marise Ph. Born

Free University of Amsterdam
The Netherlands

One of the characteristics of endemic cretinism is serious and irreversible mental retardation. Although the percentage of cretins in iodine-deficient regions varies, it may be as high as ±10% in some areas. An important question is whether the noncretinous population—the so-called healthy group—in an iodine-deficient area also exhibits some mental abnormalities, but perhaps to a lesser extent. During the past 25 years, several investigators have attempted to find an answer to this question. The outcomes of their investigations were quite divergent. Some found no correlation or no clear relationship between the presence and absence of iodine and scores obtained on an intelligence test; others established a clear relationship.

Because these studies were often small in scale, particularly as to the number of participating subjects, it seemed useful to combine and integrate their results. The traditional manner of combining the outcomes of a number of studies is to give a verbal summary. However, narrative reviews are considered to be "subjective," "scientifically unsound," and "an inefficient way to extract useful information" (1).

In a literature integration study, the simplest way to arrive at some form of quantification is the "vote-counting method." The outcomes of the various studies are classified into three groups: significant positive effects, significant negative effects, and nonsignificant effects (2). The category with the largest number of studies is then assumed to demonstrate the "true" relationship. Although vote-counting is attractive because of its simplicity, it should be discouraged as a method of research integration (1,3,4). One of the drawbacks to this method is the fact that statistical significance is highly dependent on sample size. Studies with a small effect may be highly significant if there is a large sample. For example, if each study in a set of 20 studies shows statistically significant effects, we still do not know whether the effects are large enough to be relevant.

A fairly recent and much sounder method of combining results from several studies is metaanalysis. Metaanalysis is "the quantitative synthesis of a large collection of summary statistics from individual studies on a single topic" (5). The idea behind the method is that the systematic integration of studies performed earlier is a more valuable contribution to theory formation than the addition of yet another, more or less limited, study. Metaanalytical techniques have been developed to overcome the problems of the traditional and vote-counting methods of research integration cited above.

A metaanalysis procedure generally consists of two parts:

1. combining the outcomes of various studies to arrive at a single overall estimate of the effect to be studied. In simple terms, this overall estimate is the mean of the study outcomes weighted for sample size. For example, if an investigator has 10 studies available, all of

Table 1. Information About Studies Included in the Metaanalysis

Country	n	Age (years)	Intelligence Test	r
Indonesia	102	13–20	Test Intelligensi Kollektip Indonesia	0.57
New Guinea	20	10–12	Pacific Design Construction Test	0.56
Spain	97	2.5–5	McCarthy Scales of Children's Abilities	0.53
China	499	7–14	Combined Raven's Test for Rural China	0.51
Ecuador	32	3–5	Stanford–Binet Intelligence Scale	0.50
Ecuador	46	6–10	Goodenough Draw-A-Man Test	0.48
Ecuador	67	3.5–5.5	Stanford–Binet Intelligence Scale	0.48
Spain	184	6–12	Cattell Culture Fair Intelligence Test	0.45
Ecuador	124	3.5–5.5	Stanford–Binet Intelligence Scale	0.37
Indonesia	143	6–12	Test Intelligensi Anak	0.36
Spain	74	2–30[a]	Bayley Scale of Infant Development	0.36
China	192	7–14		0.28
China	99	25–45		0.28
Indonesia	163	5–20	Raven Coloured Progressive Matrices	0.28
Zaire	138	4–23[a]	Brunet–Lézine Scale	0.26
Ecuador	50	6–10	Goodenough Draw-A-Man Test	0.24
Chile	90	±6–±12	Wechsler Intelligence Scale for Children, Revised	0.20
China	94	0–7		0.18
Ecuador	108	12–15	Wechsler Intelligence Scale for Children	0.12
Ecuador	154	8–12	Stanford–Binet Intelligence Scale	0.08
Bolivia	200	5.5–12	Stanford–Binet Intelligence Scale	0.04

[a]Age of these subjects in months.

which state the correlation between salt consumption and blood pressure, then metaanalysis is used to obtain the mean weighted correlation between salt consumption and blood pressure for the 10 studies. This mean correlation is the best estimate of the size (or strength) of the effect of salt consumption on blood pressure in a population, and is termed "the population effect size."

2. comparison of the individual outcomes, primarily focusing on the variation in outcomes. For instance, there may be large variation between the studies in the size of the correlations between salt consumption and blood pressure, in which case we would speak of heterogeneous outcomes. Small variation in outcomes would indicate a homogeneous group of studies.

The two most commonly used measures in a metaanalysis are the Pearson product-moment correlation coefficient and Cohen's d-value (6). The latter is the standardized difference in means between two groups on the same variable. Both measures indicate the size, or the strength, of an effect.

Results from individual studies are frequently reported in a variety of forms: for example, in t-values, chi-squares, and mean and standard deviations. Metaanalysis uses transformation formulas to convert these different types of indices into a correlation coefficient or a d-value. There are several statistical indicators to determine the heterogeneity of outcomes, the most important of which is that of Stoffelmayr, Dillavou, and Hunter (7).

A metaanalysis comprises a number of steps: formulation of the research issue, collecting and selecting the studies, coding characteristics of the studies, performing statistical analyses, interpreting the data, and drawing conclusions.

Research Issue

Is there a relationship between iodine presence or absence in children and adults, and cognitive development. And, if so, how strong is this relationship?

Selection of the Studies

This metaanalysis involved studies of the influence of iodine deficiency on mental development. This means that reports of studies were selected comparing iodine-deficient regions with non-iodine deficient regions, as well as publications describing the effect of iodine administration on mental development (iodine prophylaxis vs. placebo).

Because the literature on the subject of iodine and mental development is fairly limited, it was not difficult to trace the studies. A total of 21 studies proved to be suitable for the metaanalysis. (see Appendix). The data from the studies used for the metaanalysis were found in international journals, proceedings of symposia, books, IDD Newsletters, and a thesis. In principle, a study was selected if it contained information on the general cognitive functioning of children and adults living in iodine-deficient areas and if it gave the statistical data (e.g., mean and standard deviations or significance tests and sample sizes) required for the application of the metaanalytic procedure. Sometimes several studies had been done of the same sample group, using different tests, for example. In such cases, the outcomes of one study only were included. However, studies performed in the same area, but in which different intelligence tests were used for different age groups, were all selected for metaanalysis. This meant that the condition of independent data was met.

To integrate the individual studies, first of all the statistical measures (t, F, r, etc.) from the set of studies had to be converted into the same measure. The correlation coefficient was chosen as the index.

Table 1 shows a few characteristics of the 21 studies included in the metaanalysis. The studies were performed in very divergent cultures. The intellectual capacities were often determined using "Western" tests, although the tests had been adapted to the local culture in a number of cases. It is interesting to note that several

Table 2. Stem-and-Leaf Display for 21 Effect Sizes

0						
+ 0.1	4	8				
+ 0.2	2	8				
+ 0.3	0	4	6	8	8	8
+ 0.4	6	6	7			
+ 0.5	5	8	8			
+ 0.6	0	1	3	6	7	
+ 0.7						

studies used tests with a strong verbal component, which is highly sensitive to environmental and cultural influences.

The total number of individuals involved in the 21 studies was 2676, ranging from 20 for study two to 499 for study four. For each study, the total number of subjects was indicated for both the experimental and the control group. The ages of subjects ranged from 2 months to 45 years. However, nearly all studies were limited to children with a maximum age of 15 years.

As shown in the right-hand column of Table 1, all effect sizes as expressed in correlation coefficients calculated for the 21 studies were positive; that is, the noniodine-deficient group obtained a higher mean score in all studies than did the iodine-deficient group. Fourteen of these correlations were significant at the 1% level. The correlations are shown in Table 2 in a "stem and leaf" display (8). After transformation of each

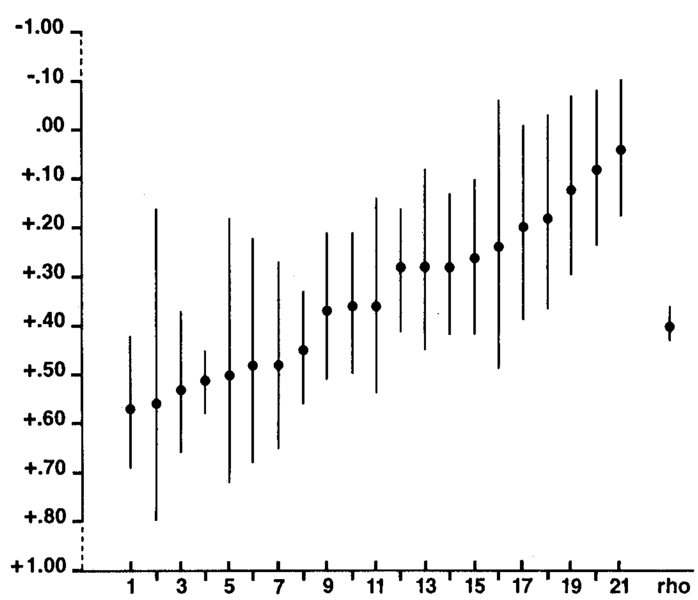

Figure 1. Visual display of effect sizes and confidence intervals for 21 studies concerning general intelligence and for the (weighted) population effect size rho (studies 19, 20, and 21 are excluded).

Table 3. Summary of Statistics Employed in Metaanalysis of Iodine Deficiency Effects

Variables	Total Group	Selected Group
Number of studies	21	18
Number of subjects	2676	2214
Population effect size (weighted mean r)	0.34	0.40
Significance (p)	0.00000	0.00000
Mean standardized difference (g)	0.78	0.90
95% Confidence interval	0.31–0.38	0.36–0.43
Test of homogeneity	heterogeneous	homogeneous

individual effect size into Fisher's z, an estimate was made of the overall effect size using the Schmidt–Hunter method (9). This method ascribes different weights to the various studies on the basis of the size of the groups studied. Studies of larger groups are ascribed more weight because these more extensive studies yield more reliable outcomes than smaller studies. The weighted population effect size calculated in terms of the correlation coefficient is 0.34 (significant at p = 0.00000). The 95% confidence interval of this population effect size lies between 0.31 and 0.38. Figure 1 gives a visual display of the effect sizes and confidence intervals.

In the interpretation of the population effect size, the degree of homogeneity of the underlying data is important. To determine the homogeneity of the studies, the measure of Stoffelmayr et al. (7) was used. This index is calculated as follows: the variance due to sampling errors is subtracted from the total variance of the group of correlation coefficients. If the remainder (the residual standard deviation) is not larger than one-quarter of the population effect size, then the set of studies may be termed homogeneous. In our case the residual standard deviation was 0.14. This is larger than one-quarter of the population effect size, which is (0.25 × 0.344) = 0.086. This means that the composition of the group of studies was heterogeneous; in other words, the studies did not come from the same population.

Closer analysis of the individual studies showed that the design of the three studies with the lowest effect sizes deviated from the other studies on an essential point. The data for these three studies had been collected from a select group consisting exclusively of school children, but no data were available on drop-outs. This was sufficient reason to reassess the homogeneity omitting these three particular studies.

The remaining 18 studies proved to be sufficiently homogeneous to assume that they came from the same population group (residual SD = 0.097, which was smaller than one-quarter of the population effect size, 0.25 × 0.40). The weighted population effect size of this reduced group of 18 studies was 0.40 (significant at p = 0.00000), with a 95% confidence interval of 0.36–0.43.

DISCUSSION AND RECOMMENDATIONS

Whereas the outcomes from a number of studies point to a negative effect of iodine deficiency on cognitive development in children and adults living in seriously iodine-deficient areas, other studies do not clearly show such an effect. The most important limitation of the studies is the often small number of subjects on which the outcomes are based. Rarely do these small studies "provide sufficiently definitive answers upon which to base policy" (10). However, metaanalysis makes it possible to integrate the outcomes of several smaller studies in a responsible fashion and to thus arrive at a weighted overall assessment, expressed in the population effect size.

A total of 21 publications could be traced that contained the statistical data required for metaanalysis. Three of these studies were omitted from further processing because the composition of the groups studied was essentially different (they were composed exclusively of school children). The remaining 18 studies formed a homogeneous group. The effect sizes of the individual studies varied from 0.12 to 0.57. The population effect size expressed in the weighted mean r was 0.40, with a 95% confidence interval lying between 0.36 and 0.43. Table 3 summarizes the data for the total and the selected group of studies.

The meaning of the outcome of the metaanalysis can be clarified on the basis of the d-value, where d is the difference between the two group means divided by the standard deviation. Co-

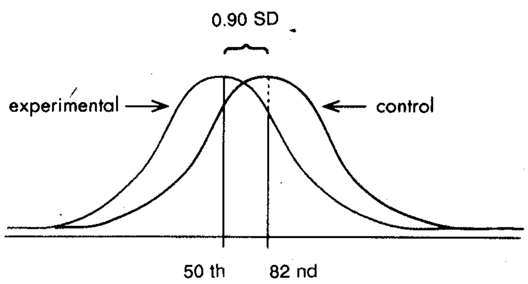

Figure 2. Average effect size in standard deviation units (50th, 82nd = percentiles of noniodine-deficient group distribution).

hen (9) classified *d*-values into small effect size (*d* = 0.2), medium effect size (*d* = 0.5) and large effect size (*d* = 0.8). In the metaanalysis of the effects of iodine deficiency on cognitive development, a large effect size was found with a *d*-value of 0.90. Concretely, this means that the mean scores for the two groups studied, the iodine-deficient group and the noniodine deficient group, are 0.9 SD, or 13.5 IQ points, apart. In other words, a typical child with an average score in the noniodine deficient group scores higher than 82% of the children from the iodine-deficient group. If the test scores of the two groups of subjects exhibit normal distributions, then the outcomes of metaanalysis can also be graphically represented in the form of overlapping distributions (see Fig. 2).

However, there is a question whether the scores of the iodine-deficient group in particular form a normal distribution. This question was also posed earlier by DeLong (11): "... whether the low mean scores in the iodine-deficient group are accounted for by a few 'sub-clinically cretinous' children, or whether iodine deficiency shifts the entire population distribution of skills to a lower level." Further study is needed of this.

Unfortunately, an unknown number of studies is missing from this metaanalysis. In the first place, not all studies could be traced, as some journals and books were not available in The Netherlands. In the second place, some publications mentioned no statistical data required to perform a metaanalysis. In the third place, there may also be studies that have not been reported, or that are described in internal publications or local journals. A subsequent metaanalytic study should include the data from as many studies as possible. Then, in addition to analysis of the general intelligence level, it would also be possible to investigate whether there are differences on more specific intelligence factors such as memory, spatial orientation, perception, etc. It would also enable the study of possible gender differences.

Many studies have been devoted to the influence of iodine deficiency on motor and psychomotor skills. A metaanalysis of the data from this group of studies will yield a more reliable picture, primarily of the possible consequences for more specific motor skills.

Call for Data

The authors urgently appeal to researchers in the field of iodine deficiency in relation to cognitive, motor, and psychomotor functioning to furnish data for the performance of a more complete metaanalytic study. The data required for metaanalysis are:

1. number of subjects per subgroup (e.g., gender, age), age mean and range, test(s) used, dependent/independent groups.
2. mean scores and standard deviations for the male and female groups as well as for the total group, or the exact data from significance tests (t, F, χ^2, etc.), stating direction of the effect and whether the test was one- or two-tailed. If more than one test was used, then the data from each test are also important.
3. if possible: length, weight, goiter incidence or frequency, goiter grade, serum T_4, serum TSH, and estimated level of iodine intake.

SUMMARY

1. Metaanalysis makes it possible to integrate the outcomes of several (smaller) studies of neuromotor and cognitive functions.
2. A total of 18 studies, which had already been performed on the effects of iodine deficiency, could be used in a metaanalysis.
3. A large effect size was found with a *d*-value of 0.90. This meant that the mean scores for the iodine- and noniodine deficient-groups were 0.9 SD, or 13.5 IQ points, apart.

REFERENCES

1. Light, R.J.; Pillemer, D.B. Summing up. The science of reviewing research. Harvard University Press, Cambridge; 1984.
2. Ma, T.; Wang, D.; Cheng, Z.P.; Wang, Y.Y. Neurointellectual impairment in Non-cretin individuals in IDD. Unpublished report; 1986.
3. Hyde, J.S. Introduction: meta-analysis and the psychology of gender. In: Hyde, J.S.; Linn, M.C., Eds. The psychology of gender: Advances through meta-analysis. Johns Hopkins, Baltimore, MD; 1986.

4. Schwarzer, R. Meta-Analysis Programs Manual. Berlin: Freie Universität Berlin, Institut für Psychologie; 1989.
5. Wolf, F.M. Meta-analysis: Quantitative methods for research synthesis. Sage Publications, Beverly Hills, CA;.1986
6. Cohen, J. Statistical power analysis for the behavioral sciences. Academic Press, New York; 1977.
7. Stoffelmayr, B.E.; Dillavou, D.; Hunter, J.E. Premorbid functioning and outcome in schizophrenia: A cumulative analysis. J. Consult. Clin. Psychol. 51:338–352; 1983.
8. Tukey, J.W. Exploratory data analysis. Addison-Wesley, Reading MA; 1977.
9. Cohen, J. Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum, Hillsdale, NJ; 1988.
10. Hedges, L.V.; Olkin, I. (1985). Statistical methods for meta-analysis. Academic Press, New York; 1985.
11. DeLong, G.R. The effect of iodine deficiency on neuromuscular development. IDD Newsletter 6(3); 1990.

APPENDIX

List of Studies Used in the Metaanalysis

Bautista, A.; Barker, P.A.; Dunn, J.T.; Sanchez, M.; Kaiser, D.L. The effects of oral iodized oil on intelligence, thyroid status, and somatic growth in school-age children from an area of endemic goiter. Am. J. Clin. Nutr. 35:127; 1982.

Bleichrodt, N. Iodine deficiency disorders in Indonesia and Spain. In Keats, D.M.; Munro, D.; Mann, L. Eds. Heterogeneity in cross-cultural psychology. Swets & Zeitlinger B.V., Amsterdam; 1990.

Bleichrodt, N.; Drenth, P.J.D.; Querido, A. (1980). Effects of iodine deficiency on mental and psychomotor abilities. Am. J. Phys. Anthrop. 53:55; 1980.

Bleichrodt, N.; Escobar del Rey, F.; Morreale de Escobar, G.; Garcia, I.; Rubio, C. Iodine deficiency, implications for mental and psychomotor development in children. In DeLong, G.R.; Robbins, J.; Condliffe, P.G. Eds. Iodine and the brain. Plenum Press, New York; 1989.

Bleichrodt, N.; Garcia, I.; Rubio, C.; Morreale de Escobar, G.; Escobar del Rey, F. Developmental disorders associated with severe iodine deficiency. In Hetzel, B.S.; Dunn, J.T.; Stanbury, J.B. Eds. The prevention and control of iodine deficiency disorders. Elsevier Science Publishers, New York; 1987: 65.

Collins, J.K.; Boyages, S.C.; Maberly, G.F.; Eastman, C.J.; Derun, Liu; Qidong, Qian; Peiying, Zhang; Chequi, Qu. The effects of iodized oil and iodized salt prophylaxis on cognitive functioning in iodine deficient endemias in China. International Conference on Iodine and the Brain, Bethesda. MD; 1988.

Djokomoeljanto, R.R.J. The effect of severe iodine deficiency: A study on a population in Central Java, Indonesia. Thesis. Indonesia, Semarang; 1974.

Dodge, P.R.; Palkes, H.; Fierro–Benitez, R.; Ramirez, I. Effect on intelligence of iodine in oil administered to young Andean children. A preliminary report. In Stanbury, J.B.; Kroc, R.L. Eds. Endemic goiter. PAHO Scientific Publication 193:378; 1969.

Fierro–Benitez, R. Iodized oil and mental development. In Medeiros-Neto, G.; Maciel, R.M.B.; Halpern, A. Eds. Iodine deficiency disorders and congenital hypothyroidism. Ache, Sao Paulo; 1986:120–126.

Fierro-Benitez, R.; Ramirez, I.; Suarez, J. Effect of iodine correction early in fetal life on intelligence quotient. A preliminary report. In Stanbury, J.B.; R.L. Kroc, R.L. Eds. Human development and the thyroid gland. Relation to endemic cretinism. Plenum Press, New York; 1972:239.

Fierro–Benitez, R.; Cazar, R.; Stanbury, J.B.; et al. Effects on schoolchildren of prophylaxis of mothers with iodized oil in an area of iodine deficiency. J. Endocrinol. Invest. 11:327; 1988.

Greene, L.S. (1973). Physical growth and development, neurological maturation and behavioral functioning in two Ecuadorian Andean communities in which goiter is endemic. Am. J. Phys. Anthrop. 38: 119–134; 1973.

Hetzel, B.S. The prevention and control of iodine deficiency disorders. United Nations: Nutrition policy discussion paper no. 3; 1988.

Fa-Fu, Lin; Aihaiti; Hong-Xin, Zhao; Jin, Lin; Ji-Youg, Jiang; Maimaiti; Aiken. The relationship of a low-iodine and high-fluoride environment to subclinical critinism in Xinjiang. IDD-Newsletter 7:3; 1991.

Muzzo, S.; Leiva, L.; Carrasco, D. Influence of a moderate iodine deficiency upon intellectual coefficients of schoolage children. In Medeiros–Neto, G.; Maciel, R.N.B.; A. Halpern, A. Eds. Iodine deficiency disorders and congenital hypothyroidism. Ache, Sao Paulo; 1986:40–45.

Pharoah, P.O.D.; Connolly, K.J.; Ekins, R.P.; Harding, A.G. Maternal thyroid hormone levels in pregnancy and the subsequent cognitive and motor performance of the children. Clin. Endocrinol., 21: 265; 1984.

Querido, A. Endemic cretinism—a continuous personal educational experience during 10 years. Postgrad. Med. J. 51:591–599; 1975.

Thilly, C.H. Psychomotor development in regions with endemic goiter. In Hetzel, B.S.; Smith, R.M. Eds. Fetal brain disorders—recent approaches to the problem of mental deficiency. Elsevier/North-Holland Biomedical Press, Amsterdam; 1981:265.

Trowbridge, F.L. (1972). Intellectual assessment in primitive societies, with a preliminary report of a study of the effects of early iodine supplementation on intelligence. In Stanbury, J.B.; Kroc, R.L. Eds. Human development and the thyroid gland. Relation to endemic cretinism. Plenum, New York; 1972: 137–149.